DOCUMENT RESUME

ED 033 043 SE 007 566

By-Chiaraviglio, Lucio

Reductionistic Inferences in Modern Biology.

Georgia Inst. of Tech., Atlanta. School of Information Sciences.
Spons Agency-National Science Foundation, Washington, D.C.
Report No-GITIS-69-05
Pub Date 69
Note-32p.
EDRS Price MF-$0.25 HC-$1.70
Descriptors-Biology, *Genetics, *Logic, *Philosophy, Science History, Scientific Methodology, Validity

The author analyzes the logic of inferences in modern biology which serve as reductionist bridges between the organismic and molecular levels of explanation. He distinguishes pragmatic validity from deductive or inductive validity, and discusses the requirements for validity of pragmatic inferences (which are accepted or rejected in terms of the performances of their premises and conclusions). The development of reductionist bridges between classical and molecular genetics is discussed, using the genetic role of nucleicacids and the co-linearity of genetic and physical structures as examples. The experimental procedures and results that led to inferences in these areas are described and analyzed in terms of the performances of sets of sentences. It is shown that the validity of such inferences is dependent on the compatability of conditions of performances of the sentences. The importance of collateral knowledge in selecting compatable conditions of performance is illustrated. (EB)

DOCUMENT RESUME

ED 033 043 SE 007 566

By-Chiaraviglio, Lucio

Reductionistic Inferences in Modern Biology.

Georgia Inst. of Tech., Atlanta. School of Information Sciences.
Spons Agency-National Science Foundation, Washington, D.C.
Report No-GITIS-69-05
Pub Date 69
Note-32p.
EDRS Price MF-$0.25 HC-$1.70
Descriptors-Biology, *Genetics, *Logic, *Philosophy, Science History, Scientific Methodology, Validity

The author analyzes the logic of inferences in modern biology which serve as reductionist bridges between the organismic and molecular levels of explanation. He distinguishes pragmatic validity from deductive or inductive validity, and discusses the requirements for validity of pragmatic inferences (which are accepted or rejected in terms of the performances of their premises and conclusions). The development of reductionist bridges between classical and molecular genetics is discussed, using the genetic role of nucleicacids and the co-linearity of genetic and physical structures as examples. The experimental procedures and results that led to inferences in these areas are described and analyzed in terms of the performances of sets of sentences. It is shown that the validity of such inferences is dependent on the compatability of conditions of performances of the sentences. The importance of collateral knowledge in selecting compatable conditions of performance is illustrated. (EB)

Work reported in this publication was performed at the School of Information Science, Georgia Institute of Technology. The primary support of this work came from the Georgia Institute of Technology and from agencies acknowledged in the report.

* * *

The graduate School of Information Science of the Georgia Institute of Technology offers comprehensive programs of education, research and service in the information, computer and systems sciences. As part of its research activities the School operates, under a grant from the National Science Foundation, an interdisciplinary science information research center. Correspondence concerning the programs and activities of the School may be addressed to Director, School of Information Science, Georgia Institute of Technology, Atlanta, Georgia 30332.

Telephone: (404) 873-4211

GITIS-69-05

(Research Report)

REDUCTIONISTIC INFERENCES IN MODERN BIOLOGY

Lucio Chiaraviglio

School of Information Science 1969

GEORGIA INSTITUTE OF TECHNOLOGY

Atlanta, Georgia

ACKNOWLEDGEMENT

The work reported in the paper has been sponsored in part by the National Science Foundation Grant GN-655. This assistance is gratefully acknowledged.

VLADIMIR SLAMECKA
Project Director

CONTENTS

INTRODUCTION

We shall explicate a notion of pragmatically valid inference that is aimed at providing a reconstruction of the manner in which certain key inferences have been constructed and accepted in modern reductionistic biology. The examples of inferences that we shall consider span classical and molecular genetics. The principal aim of these inferences is to correlate the phenotypic and genotypic properties, functions, and structures of organisms to their molecular machinery. The inferences are reductionistic bridges between the organic and molecular levels of explanation. They are reductionistic bridges in the sense that these inferences constitute grounds for selecting among alternative explanations of the causal connections between the molecular machineries and the classical-genetic features of organisms.

The inferences that concern us are pragmatic. They are accepted or rejected in terms of the performances of their premises and conclusions. As a matter of presentational simplification we shall consider that performances settle the acceptance or rejection of sentences rather than the degree of their acceptance or rejection. But it is also true that the sentences of most of the inferences that concern us are accepted or rejected on the strength of the relevant assays. Assays that settle degree of acceptance are relatively rare. Thus we may talk without too much distortion of assays that are germane to the acceptance or rejection of sentences. Such assays will be called the condition of performance that are germane to a sentence.

If we let S_o be a sentence of classical genetics and S_m a sentence of molecular genetics, then we may say that an inference from S_o to S_m (or Mutatis mutandi an inference from S_m to S_o) is pragmatically valid if and only if: a) there is a rule R that states that C_{o1}, C_{o2}, ..., C_{on} are conditions under which S_o is to be performed, that C_{m1}, C_{m2}, ..., C_{mn} are conditions under which S_m is to be performed, and that for any C_{oi} under which S_o is performed, there exists a corresponding C_{mi} under which S_m is to be performed; b) every pair of conditions taken from the above sets are

compatible; c) to every performance in which S_o is accepted R never correlates a performance in which S_m is rejected; and d) there is a condition under which S_o is accepted.

R is a rule that regulates the conduct of performance. We shall see in the context of specific inferences various examples of such rules. We do not claim that in fact such rules are usually stated in the literature. What does occur in the literature under the heading of "materials and methods" are descriptions of the performances. Furthermore, such descriptions usually give the details of only those aspects of the performances that depart from well established experimental practice.

It is clear that pragmatic validity depends on the conditions of performance that are available or have been devised at the time of performance. So it can happen that inferences that are considered to be valid at one time turn out to be invalid or fail to be pragmatic inferences entirely at another time. Experimental methods may evolve in such a way that some assays that were considered germane at one time are seen to be inadequate. At such a juncture we may say with equal justice that either that the rule R has been scrapped or that the rule is retained but that the conditions of performance which R correlates are changed. In the first case we conceive of R to be stated relative to a specific set of assays. In the second case we conceive of R as a rule that may govern performance relative to a type or kind of assay. Either conception will suit our purposes. What is essential to our discussion is that R be understood to regulate performance. That is to say that R must be understood to be not merely a description of the correlation among assays that are germane to certain sentences, but must be also understood as an imperative commanding certain actions. The actions commanded are the performances of the sentences in question.

The performances are appropriate to the settling of the pragmatic validity of some inferences. One of the conditions that have to be fulfilled in order to achieve this end is that the rule of performance only correlates pairs of conditions that are compatible. We may say that two conditions of performance C_1 and C_2 are compatible relative to a set of

performable sentences S if and only if there is no sentence s in S such that s is rejected on condition C_1 and accepted on condition C_2. If C is a set of conditions which are germane to the performances of the sentences in S, then compatibility is a similarity relation on C. Requirement b), the requirement of compatibility, states that the rules R that correlate performances must be selected from among those that induce on the set of conditions C similarity relations that are subrelations of compatibility on C. This requirement makes it possible to partition a set of performable sentences S relative to a set of germane conditions C into pragmatically equivalent classes of sentences. This can be done only if the R that generates such a partition meets the requirement of compatibility. Of course we assume that any notion of pragmatic inference that is reasonably relative to a given set of performable sentences should yield an equivalence partition of this set.

We may note in passing that entirely similar restrictions on the rules of performance as the one just explicated may be derived from compatibility requirements on the set of performers, the set of times of performance, and other sets of factors on which performances may depend. This point has been elucidated elsewhere (Chiaraviglio, 1969).

Requirement d) is a fairly straight forward pragmatic rendering of the truth requirement for deductive inferences. The last requirement bars the pragmatic acceptance or rejection of inferences whose premises are rejected under all the conditions that are germane to the premises. It is our impression that inferences whose premises are rejected are not accepted or rejected on pragmatic grounds.

Since the rule R governs conduct, we might say that pragmatic inferences are deontological. It might be thought that the rules that govern the conduct of performance may be taken as inferential principles that are parallel to material principles of inference. But there is a big difference between the two cases since in the material inference case the inference that is obtained by conjoining the material principle to the premises is a deductively or inductively valid inference while in the case of

pragmatic inferences the addition of the rule to the premises will not yield inferences that are inductively or deductively valid. The passage from pragmatically valid inferences to deductively or inductively valid inferences is quite complex since the subject matters of the premises, conclusions, and rules are heterogeneous.

What seems to happen is that given a rule R that is adequate, then the validity of some of the pragmatic inferences that R allows is taken as evidence for or against some biological hypotheses that could function as material principles of inference for the set of sentences in question. In other words, the pragmatic validity of a set of inferences is used as one of the criteria of selection among alternative explanatory proposals. The examples that follow will illustrate how this is sometimes done and what sort of warrants are proposed for the adequacy of the rules.

EXAMPLES

Part of the success which molecular genetics seems to enjoy is in great measure due to the acceptance of inferences of the type that concern us. The feeling that they are well founded and that they constitute good evidence in favor of some of the principal tenents of molecular genetics has led to a rather rapid acceptance of some of the molecular reconstruction of classical genetic concepts. This is true even in those areas of biology where direct evidence that bears on the adequacy of the molecular reconstruction is not presently available. For example, the proposal that a gene is a set of contiguous loci of bases in the polymers that are the genetic materials has gained part of its acceptance through a large set of inferences that relate classical genetic structures, map structures, to the structures of these polymers. Most often the evidence for such inferences comes from viral and bacterial systems. Nevertheless, it is generally assumed that the correlation between genetic and molecular structures in higher organisms is of the same type. The proposed molecular reconstruction of "gene" is taken to be appropriate to all of biology and not just a parochial feature of bacterial and viral biology.

The fact that some of the inferences that concern us occur rather

early in the development of molecular genetics may not be a matter of mere historical accident. In many ways these inferences seem to be required early steps in the methodology of reduction. Of course, it goes without saying that none of the possible bridges between the molecular and classical levels would have been thought of if the way had not been paved by purely molecular and purely genetic discoveries. The discovery and elucidation of some of the principal structures of macromolecules such as proteins, DNA, and RNA seems to have been obligatory steps in the reductionistic development. Similarly the elucidation of the genetics of organisms which reproduce rapidly and thus are well suited to fine structure exploration of genetic maps was probably also required. The rapidly reproducing organisms often used, viral and bacterial systems, happened to lend themselves to biochemical procedures that led to the in vitro reassembly of some of the principal molecular components and processes. These and many other factors probably played important roles in the development of reductionistic hypotheses.

The detailed historical task necessary to assess the various roles played by such factors is beyond the scope of this paper. Nevertheless, we hope that much of what we say or presuppose should be amenable to confirmation or falsification through a careful study of the relevant history.

The Genetic Role of Nucleic Acids

We may now outline briefly a number of experiments that were instrumental in showing that certain nucleic acid components of virus and bacteria were the genetic materials of these organisms.

It had been known for some time that nucleic acids were the principal components of nuclei of cells. As early as 1944 it had been shown that in bacterial transformation the hereditary factors of the donor bacillus were transferred to the genome of the receiver bacillus through the exclusive action of the donor DNA molecules (Avery, MacLeod and McCarty, 1944). In 1952 it was shown that T2 phage DNA is the principal component that gains entrance into the host cell upon infection and that the DNA that gained

entrance was capable of producing viable offsprings without the further collaboration of the remaining T2 components (Hershey and Chase, 1952). In 1956 it was shown that Tobacco Mosaic Virus (TMV) that was artificially reconstituted from proteins and nucleic acids extracted from two distinct strains which differed in their coat proteins and host symptoms produced offsprings whose type was identical with the type of the TMV from which the nucleic acids were obtained. Furthermore, it was also shown that the extracted nucleic acids by themselves were sufficient to produce infection and viable TMV progeny (Fraenkel-Conrat, 1956; Gierer and Schramm, 1956). By 1962 similar evidence as to the genetic role of nucleic acids had accumulated for a number of different virus (Cold Spring Harbor Symp. Quant. Biol., 1962).

In broad outlines, the procedures that were followed in order to correlate genetic functions with some of the molecular components of organisms may be summarized as follows.

First, there is an identification of portions of the life cycle of the organism during which certain molecular components may be manipulated so as to effect some change or a lack of change in the genetic make-up of the progeny that is the outcome of the cycle. For example, in the case of bacterial transformation it was noted that by injecting mice with a large amount of heat killed Type III pneumococci and a small living innoculum of Type II R one could recover from such mice living Type III pneumococci (Griffith, 1928). Later it was shown that the same effect could be obtained in vitro (Alloway, 1932). An infection with Type II R could be made to yield of Type III if it occurred in the presence of heat killed Type III pneumococci. The genotype of the offspring could be manipulated by means of some element derived from heat killed bacteria. In the case of phage we have the identification of a life cycle which consists of an attachment of the virus particles to the host cell, a period in which there is an absence of intra-cellular infective units (the eclipse period) and a period in which infective units accumulate in the host until it bursts and liberates a multiplicity of infective virus particles (Wollman and Wollman,

1936; Doermann, 1948). The identification of the eclipse period disclosed the possibility that the infective unit was composed of several components that might be separable physically and whose functions in viral replication might be thereby sorted. Similar remarks may be made about TMV whose life cycle of infection, replication and release has a first stage in which one of the TMV components is left outside the host.

Second, there is an identification of some of the molecular components of the organisms in question. For example, crude extracts of heat killed pneumococci were known to contain polysaccharides, somatic carbohydrates, nucleo proteins, and free nucleic acids of various types. In the case of phage and TMV the principal or only components were found to be coat proteins and encased nucleic acids.

Third, there is an identification of one molecular component which is indispensable for the production of the observed effect. Of course, such an identification is successful only if the indispensable element is one of the previously identified components. For example, in the case of transformation, one of the nucleic acids found in crude extracts of heat killed Type III pneumococci was responsible for the observed genetic transformation. In the case of TMV the RNA encased in the protein coat was by itself capable of producing infection of tobacco leafs and giving rise to what appeared to be fully competent TMV progeny. In the case of phage T2 the protein coat component could be separated from the infected cells which contained the T2 DNA component. These infected cells were capable of producing fully viable T2 progeny without the further intervention of T2 coat proteins.

In all three cases the nucleic acids carried either physical labels, genetic labels or both. The DNA involved in the transformation experiments was obtained from pneumococci of a genetic type distinguishable from the pneumococci that it transforms. The RNA used in the artificially reconstituted TMV experiments was obtained from viruses of a genotype that was different from the genotype of the viruses from which the coat proteins were obtained. The phenotypic differences that scored the genotypes were

in these cases both a difference in coat proteins and host symptoms. In the T2 experiments two stocks of standard T2 were radioactively labeled. One stock carried the radioactive label in its proteinaceous component. The other stock carried the radioactive label in the nucleic acid component. The recovery of the genetic label in the offsprings produced in the cases of pneumococci transformation and reconstituted TMV was taken as evidence for the genetic role of the DNA and RNA components respectively. The presence of the radioactively labeled T2 DNA in the infected cell and the fact that the radioactively labeled T2 protein components could be stripped from the cells without altering the subsequent production of T2 progeny was taken as evidence of the genetic role of T2 DNA. Thus, the ability of genetically or physically labeled material to produce the offsprings of the appropriate type is evidence for the genetic role of the material.

The pragmatic inferences performed in the three cases sketched involve the performance of three distinguishable sets of sentences. The first set may be called the "organismic functions descriptions." These sentences specify relevant portions of the life cycle of the organism and phenotypic and genotypic properties that are characteristic of the organism in question (e.g., transformation, infectivity, capability of producing offsprings, capability of producing offsprings of a given type, etc.). The second set of sentences may be called "molecular components identifications." These sentences identify some or perhaps all of the molecular components or types of molecular components at some portion of the life cycle of the organism. For example, free TMV is composed of proteins and nucleic acids and these types are represented in each particle by a single RNA molecule and a large number of identical protein molecules. The third set of sentences may be called "molecular functions descriptions." These sentences ascribe to some of the identified molecular components some organismic functions. For example, the transforming function of heat killed Type III pneumococci is ascribed to certain DNA molecules found in the crude extracts, the ability to produce viable T2 progeny is ascribed to the nucleic acid components of T2, and the ability to produce viable TMV

and the ability to infect is ascribed to the RNA component of the virus.

Language that refers to molecular components occurs in the last two sets and not in the first. The inferences that concern us have as premises the conjunction of sentences taken from the first and second set and have as conclusion sentences from the third set. Schematically: (organismic functions descriptions) and (molecular components identifications) → (molecular functions descriptions). It is clear that in order to construct deductively or inductively valid inferences on this pattern one must be able to add premises which affect a correlation between the molecular and organismic vocabulary. From present hindsight one might say that in the case of T2 one might add a premise that specifies how the T2 DNA replicates, is transcribed and is translated with the aid of the host's molecular factories and how these processes lead to intra-cellular accumulation of T2 components which are assembled into the finished product. At the time at which the pragmatic inferences were proposed very little knowledge of these processes existed. Thus any proposal of explanatory premises would have been ad hoc and highly speculative. Indeed, at the time the acceptability of the pragmatic inferences were the strongest grounds available for discriminating between alternative speculation about the molecular basis of some of the known organismic functions.

The premises and conclusions of the pragmatic inferences were performed and accepted. Furthermore, the experimental methodology for the assays that were the condition of performance of the premises and conclusions suggested a rather natural correlation between these conditions. Thus, we may describe the general form of the rules R that correlate performances in the three cases discussed as follows: First, perform organismic functions descriptions with a set of populations and select from this set a subset of populations that are uniformly labeled and are at the appropriate state of their life cycle (e.g., two radioactively labeled populations of free T2 viral particles, encapsuled pneumococci populations of Type III and Type IIR, populations of free TMV of known coat and host symptoms phenotype). Second, perform molecular component identifications

with aliquots of the relevant populations selected in the first step. Third, perform molecular functions descriptions with extracts of the populations or with the populations whose components were identified in the second step.

The correlation between the conditions of performance of organismic functions descriptions, molecular component identifications, and molecular functions description may have appeared rather natural to the investigators. All of them were well acquainted with the pattern of correlations between performances generally followed in biochemical studies of the biological activity of extracted and purified molecular components. For example, the isolation and purification of enzymes and enzyme activity exhibits a similar correlation of performances.

Of course, apparent naturalness and precedents are not a sure guide to the adequacy of the rules of performance. Furthermore, in the two viral cases extraction and purification of the nucleic acid component did not enhance the originally observed effect, the production of infection and subsequent release of progeny. Thus a simple test of adequacy based on the enhancement of the original effect would not have worked. TMV RNA by itself is less infective than the complete virus and infection of E. Coli by purified T2 DNA has not proven possible. In these instances conditions had to be found under which the functions of the two components could be isolated and correlated to the terminal effect. Thus the adequacy tests for the rules would have to provide independent checks for separation of the functions of the component molecules. In this connection it was fortunate that virus pass through a stage in their life cycle in which their nucleic acids are the only or almost the only material links connecting one generation to the next. The earliest and rather indirect evidence for this fact was evidence of the occurrence of a period in the life cycle of the virus in which no infective units were to be found. Something not fully competent as a virus but which lead to virus progeny was present in this eclipse period.

The first experiments dealing with the transfer of parental material

to the offsprings occurred in 1948 (Putnam and Kozloff, 1950). In 1950 it had been shown that the DNA of superinfecting T2 particles, unlike the DNA of the first particles to infect, appears quickly in the culture medium in acid soluble form (Lesley, et al., 1950). This suggested that the attachment of particles to bacteria is followed by a separation of phage components which exposes the DNA. The suggestion was confirmed in 1951 by allowing phage to attach to heat killed bacteria in the presence of deoxyribonucleous (Graham, 1953). In experiments reported at the Cold Spring Harbor phage meetings of the summer of 1951, Watson and Maaløe had noticed that much of the labeled parental protein in the lysates remained attached to the cellular debris and that very little of it seemed to be transferred to the progeny.

Hershey (1966) claims that evidence of this sort played an important role in his conception of the T2 experiments mentioned. It is also clear that such evidence, while in no way a definitive test of the adequacy of the rule of performance followed, constitutes historically prior grounds for the rule. Indeed, all evidence that testifies to the functional separability of the T2 components strengthens the confidence that may be placed on the rule. This is so because the rule commands a set of correlated performances which are appropriate and can be carried out only if the T2 components can be separated. If the T2 components cannot be functionally separated, then the inference cannot be constructed. If the T2 components can be functionally separated, then the inference may be constructed even though it may turn out to be invalid.

As late as 1955, there were researchers that proposed that the transforming activity of the heat killed Type III pneumococci was due to a protein component even though as early as 1949 there existed active preparations whose protein content was at most 0.02% (Hotchkiss, 1966). The proponents of these alternative explanations of the transformation phenomena considered that the pragmatic inference proposed by Avery, et al. was invalid. The presence of something less than 0.02% protein in an active DNA preparation used in the performances of molecular function descriptions led them to the

non-acceptance of these descriptions.

Of course, it was not the mere presence of the protein impurity that lead to such conclusions. There were well established hypotheses that seemed to entail an active role for protein and no similar hypotheses for DNA. Hotchkiss (1966) describes the intellectual climate for noncredence with respect to DNA activity as follows:

> "Sometimes the so-called "simplest" (and as it turned out correct) hypothesis is the offspring of a wedding of enthusiasm with ignorance. If genetically active DNA was heresy, it was accompanied by evidence, and was opposed by ideas that had become ingrained with little evidence (heresy versus hearsay, almost). Avery's own earlier career had centered considerably around the demonstration, against prevailing doctrine, that immunological specificity of antibodies could be sometimes directed towards polyesacharide as well as toward protein antigens. But the antibody is of protein nature, and all textbooks of the 30's and 40's told one that biological catalysts are proteins, and that nucleic acids are obscure but repetitious assemblages of sugar, nitrogen bases, and phosphate. How could something as specific as a transforming agent for antigen type be other than a protein? Many knew of the pitfalls - of the surprises and embarrassment good chemists have experienced at various times when an enzyme they have been pursuing is suddenly purified so much that its solution seems to be "protein free" to ordinary tests, although the activity is higher than ever. A non-protein form of pepsin had once been reported by Willstatter, for example."

To sum up then, the adequacy of the rule of inference is based on evidence for the assumption that the components in question can be functionally separated. The rule of inference may be found adequate but yet the inference itself be invalid. In such a case the rule correlates the performances appropriately and thus also correlates the conditions of

performance appropriately but the performance of the conclusions under the correlated conditions yields the nonacceptance of the conclusions. This was the position taken by the proponents of the hypothesis that protein was the transforming agent in the case of pneumococcal transformation. These researchers thought that the rule was adequate since they did not dispute the correlation between the condition of acceptance of premises and conclusions but rather argued that the presence of protein in the active preparation was sufficient to explain the observed effect.

The Co-linearity of Genetic and Physical Structures

In 1953 Watson and Crick published their epochal papers on the structure of DNA and what they perceived to be the principal genetic implication of this structure. The DNA molecules that are the genetic materials of many organisms are double stranded polymers constituted by two helically intertwined polynucleotide chains. Once it had been discovered that the DNA of certain organisms was their germinal substance, then it became pertinent to ask about the relation between genetic structures and the physical structure of the DNA.

The discovery of genetic recombination among viruses occurred in the late forties (Delbruck and Bailey, 1946; Hershey, 1946). At the same time the phenomenon was also discovered in bacteria (Lederberg and Tatum, 1946). These discoveries made possible the extension and refinement of classical genetic modes of analysis to these organisms. By 1953, Visconti and Delbruck had formulated a theory of recombination for bacteriophages that was capable of accounting quantitatively for the recombination frequencies produced under various experimental conditions (Visconti and Delbruck, 1953). The partial one way transfer of genetic factors from a donor to a recipient bacterium in E. coli was first reported in 1952 (Hayes, 1952). These and other developments made possible the construction of genetic maps of many bacteria and bacteriophages.

The genetic map is an ordering of the genetic factors of an organism that is obtained via recombination experiments. In bacteria and phages

such maps exhibit non-branching linear structure which is sometimes circular. In other words it is possible to seriate by means of recombination experiments a set of closely linked genetic factors. This series is unique or unique up to a circular permutation.

Since the genetic materials had turned out to be linear polymers, double or single stranded sequences of bases, the situation suggested that there might be a simple colinearity relation between the sequence of genetic factors and the sequence of bases of the polymers. Our present examples of pragmatic inferences are taken from a set of inferences that were instrumental in confirming this suggestion.

In order to test for the colinearity of genetic map and sequence of bases in the DNA molecule one might try to isolate genetically active DNA from an organism of known genetic constitution and then try to introduce different portions of this DNA into a situation in which there is an integration of these DNA fragments into the genetic materials of a population of organisms. By inspecting the genetic make-up of this resulting population, one might ascertain the order in which the DNA fragments incorporated carry the genetic factors of the organisms from which it was extracted. Since it is possible to construct a map, a serial ordering, of the genetic factors of the organism that is the donor of the DNA, then comparisons between the order of the genetic factors of the donor with the order in which these factors reappear in the recipient population together with the known mode of fragmentation of the DNA will allow comparison between genetic and DNA structure. In outline, this is what was done both in the case of the phage λ and in the case of E. coli.

In 1960, it was shown that genes carried by a free DNA extracted from a phage λ with known genetic constitution could be recovered in the progeny of a mixed infection with intact phage λ and the free DNA extract (Kaiser and Hodgness, 1960). It had been shown that the application of hydrodynamic shear breaks DNA molecules at approximately their mid point (Hershey and Burgi, 1960). These two facts were used to construct a test of the hypothesis of colinearity of genetic factors and DNA sequence as follows: a)

population of λ was prepared that carried three genetic markers linked in the sequence A,B,C, A and C where distal markers and B a central marker closer to A than to C in the genetic map of λ ; b) the DNA of this population was extracted and purified; c) the DNA extracted was subject to the appropriate hydrodynamic shear and broken into half molecules; d) a mixed infection with half molecules and intact λ was carried out at a concentration so low that each bacterium would receive no more than one-half molecule; and e) the recovery of the markers A,B, and C was noted in the progeny that issued from the mixed infection. The results of the last step were that A and B were found associated in the progeny while C was never found associated with A. This result would indicate that the λ DNA molecule carries the genes A,B and C in the same order as the genetic map of λ. The genes carried by half molecules are halves of the recombination map of λ. (Kaiser, 1962; Radding and Kaiser, 1963). In order to infer the physical sequence of markers in the half molecules a great number of breaks was necessary. This was done and the results confirmed the colinearity of map markers and sequence of DNA at a higher level of resolution (Kaiser and Inman, Inman, 1965).

Since in E. coli there exists a conjugation system in which a donor bacterium transfers its genetic factors to a recipient bacterium, it was possible to test for the colinearity of the genetic map of E. coli with the physical sequence in the DNA by means of experiments that made use of the naturally occurring conjugation system.

Dating from the discovery of the high frequency of recombination, Hfr, strain of E. coli in 1950 (Cavalli-Sforza, 1950), the general picture of the mating phenomena that emerged was as follows: When Hfr and F^- bacteria are mixed together specific pairings between the donor bacterium which is Hfr and the recipient bacterium F^- are rapidly made. A conjugation tube is formed between the members of the pairs and one of the chromosomes of the donor is transferred to the recipient bacterium through this tube. The transfer starts always at the same end, O, of the donor chromosome. The migration of the chromosome occurs at an apparently constant slow rate.

It is calculated that the transfer of the chromosome from Hfr to F⁻ occurs in about two hours. In the course of the transfer spontaneous ruptures of the chromosome may take place. The probability that a genetic character of the donor bacterium be transferred to the recipient decreases with the distance of the character from a certain origin O of the donor chromosome. The characters that are more distant from O are transmitted less frequently than the characters nearer to O. The notion of distance that is germane here is obtained by recombination experiments which due to the peculiarities of the E. coli mating system are quite complex. The details of these experiments need not detain us here (see Jacob and Wollman, 1961). For our purposes, it is sufficient to remark that the slow and oriented transfer of donor genes to recipient bacteria is scored by noting the recombinant frequencies for these genes in the offsprings that ensue from the matings.

The occurrence of a slow transfer process made possible experiments which would physically break the DNA being transferred at various times after the initiation of the transfer. The naturally occurring gradient of transmission of genetic characters probably suggested that such experiments could yield evidence for or against the hypothesis that the ordering of the genetic characters in a recombination map was colinear to the sequence of bases of the DNA.

Hayes was one of the first investigators to study the kinetics of the transfer of genetic characters (Hayes, 1955, 1957). What he did was to mate Hfr bacteria which carried a number of genetic characters whose order was known with F⁻ bacteria that were resistant to the virulent phage T6. These recipient bacteria carried the alleles of the selected genes. At various times after mating had occurred the donor bacteria were destroyed by adding T6 phage and the progeny of the unaffected recipient bacteria were scored for the presence of recombinants involving the selected genetic characters. Essentially the same results were obtained by experiments which separated the mating pairs at various times after mating had occurred by means of the violent agitation of the medium with a blendor (Wollman and Jacob, 1955,

1958). Both types of experiments showed that the linear order of the genetic characters of the donor was colinear with the order of the times at which mating was interrupted. Furthermore, the temporal gradient of character transmission and the genetic distances between characters were proportional.

Another method used was to grow the donors in a medium containing radiophosphorus. The radiophosphorus will incorporate in the DNA of such organisms. Thus this method introduced a physical marker in the DNA of the donors. This made possible the concurrent monitoring of physical and genetic labels.

Various lines of evidence showed that the principal effects of radioactive phosphorus decay in bacteria are caused by the decay of those atoms that have been incorporated in the bacterial DNA. The disintegration of a radiophosphorous atom in the DNA usually ruptures one of the strands of the molecule. But in a certain fraction of the disintegrations the energy liberated is large enough so as to rupture both strands. Such events break the molecule into two parts.

Radiophosphorus that has been integrated into the DNA of Hfr bacteria may be allowed to disintegrate for varying lengths of time before conjugation as well as after conjugation. When disintegration is allowed to proceed for varying lengths of time before conjugation the ruptures produced in the DNA has genetic consequences that are entirely similar to those produced by the mechanical or phage interruptions of mating. In other words, the more distant a genetic marker is from the origin O of the Hfr chromosome, the faster is its rate of elimination. The radiophosphorus decay that occurs in Hfr donors prior to mating affects the transfer of chromosonal determinants in a way that is proportional to the number of phosphorus atoms that lie between the locus of O and the locus of the determinant. On the other hand, if radiophosphorus decay is allowed to proceed after mating the rate of inactivation of different genetic markers is not dependent on their position on the chromosome (Fuerst, Jacob, and Wollman, 1956, 1961; Jacob and Wollman, 1958).

The genetic markers of Hfr bacteria can be arranged in the same linear order on the basis of the time of entry into the recipient bacteria as measured by the times of mating interruption and on the basis of the sensitivity of their transfer to radiophosphorus decay. Furthermore, this linear order is the same as that obtained by recombinational analysis. The agreement between these three modes of seriating markers is not only topological but also to a surprising degree quantitative. The relative distance of markers in terms of times, in term of sensitivity to radioactive decay and in terms of recombination distances may be made to agree by selecting appropriate constants of proportionality (Fuerst, Jacob, and Wollman, 1956, 1961; Stent and Fuerst, 1960).

The experiments just described were taken as good evidence for the hypothesis that the ordering of genetic factors in the recombination map and the ordering of the loci of these factors on the DNA molecule were colinear. The pragmatic inferences involved the performances of three distinguishable sets of sentences. The first set may be called "the genetic map descriptions" of the organisms in question. These sentences specify a set or sets of genetic markers, their order, and the recombination distances between them. The second set of sentences may be called "DNA loci descriptions." In the cases here discussed the loci in question are unbroken segments of the DNA molecule that are ordered. In the case of λ the order was obtained by halving the DNA, then halving again these halves and so on. In the case of the E. coli experiments the ordering of the segments was obtained by breakage of the DNA at various intervals from a fixed origin. The loci description sentences specify these orderings of the loci. The last set of sentences may be called "order comparisons" and they specify what genetic markers should be found associated with what DNA loci if the maps and DNA loci are to be colinear.

We may schematize the pragmatic inferences as follows: (the genetic map descriptions) and (DNA loci descriptions) → (order comparisons). The genetic map descriptions were performed and accepted by means of recombination experiments. The DNA loci descriptions were performed and accepted

by means of experiments involving molecular shear in agitated fluids or molecular breakage by radiation, and finally the order comparison sentences were performed and accepted by experiments that consisted of producing offsprings whose DNA molecules were recombinants formed from known fragments of DNA from a genetically known parent and full DNA molecules from another parent whose relevant genetic make-up was also known.

The procedure followed here consisted in taking aliquots of populations of DNA donors and recipient parents for whom the genetic map descriptions had been performed. These two parental populations were allowed to mate in the case of E. coli and the DNA transfer was interrupted at various times by the means described. Or in the case of λ the parental donor population was subjected to DNA extraction and fragmentation procedures and mixed infections were made with intact λ and DNA fragments. In either case, there was a performance of DNA loci descriptions. The offsprings that issued from these interrupted mating and mixed infections were examined for their genetic constitution and thus a comparison of the map order and the physical order of the loci was made.

As before, we may summarize the role R that correlates the performances as follows: First, perform map descriptions for a set of organisms and select from this set two parental populations such that if one parent carries the genetic marker A^+ then the other parent carries the allele A^-. The populations must be selected so that the markers are sufficiently spread over the genome and are numerous enough so as to test their order relative to the fragmentation procedure that is countenanced. Second, perform DNA loci descriptions with the DNA of one of the parental populations (the donor parent in the case of E. coli). Third, perform order comparisons by noting the recovery of genetic markers in the offsprings of the matings or mixed infections.

The tests that were offered for the adequacy of the described procedures consisted of the following: a) experiments that showed that the fragmentation procedure did fragment the DNA molecules in the manner expected (Hershey and Burgi, 1960; Stent and Fuerst, 1960; Fuerst and

tent, 1956; Hershey, Kamen, Kennedy and Gest, 1951; Stent, 1953); b) experiments that showed that the fragmentation procedure did not alter the viability of the organisms or molecules, or that it altered their viability in a predictable fashion (Anderson, 1949; Hershey and Chase, 1952); c) experiments that showed that the fragmentation procedure used produced fragments which were subsequently incorporated in the offsprings (Hodgness and Simmons, 1964); Kaiser, 1962; Radding and Kaiser, 1963; Garen and Skaar, 1958; Fuerst, Jacob and Wollman, 1956, 1961; Jacob and Wollman, 1958); and d) experiments that showed that the fragmentation procedure used did not by itself alter the genetic constitution of the progeny which ensured (e.g., the tests described which showed that the radiophosphorus decay that occurs after mating has effects on the recovery of genetic factors that is entirely independent of the map position of the factors). These tests would be evidence for the adequacy of the rule that correlates performance.

It must be noted that the rule R also depends on the truth of a number of hypotheses which were not performed during the course of the investigations regarding colinearity. Some of the hypotheses that were assumed to be true in order to construct the experiments in question were the Watson-Crick DNA structure descriptions, the genetic role of DNA, and the non-genetic roles of other molecules. It would have been indeed odd, but not in principle impossible, to obtain the results described if some or all of the above hypotheses were incorrect. Nevertheless, it is quite clear that these hypotheses provided part of the motivation for the construction of the experiments but more importantly they acted as collateral knowledge in the correlation of experimental procedures.

CONCLUSIONS

We have mentioned in the introduction that one of the conditions that a rule of performance R must meet is that it correlates only conditions of performance that are compatible. Compatibility is a similarity relation on conditions of performance that depends on a set S of performable sentences. As we have seen in the above examples there is in practice no

prior identification of a set S of performable sentences in terms of which a relation of compatibility among conditions of performance may be identified. On the contrary, all the investigations described started out with a very small set of sentences to be performed and a wealth of specific knowledge about biological matters and experimental procedures employed in the past. Somehow out of this material there condensed some fairly definite proposals as to what could be done in order to perform the sentences in question. Following Quine we will call all of this wealth of knowledge the collateral knowledge for performance (Quine, 1960).

Collateral knowledge may be explicit. That is, an investigation may be able to state the biological hypothesis or describe the conditions that experimental practices must meet in order to be bona fide tests of a given hypothesis. Collateral knowledge may also be implicitly exhibited in the correlations among performances of a given hypothesis. In both cases we shall speak of collateral knowledge since it often true in biology that there exist indices that testify to the attainment of correct results without there being any satisfactory explanation of the underlying apparati and practices. Perhaps the explicitness of collateral knowledge is a matter of degree. For us the important fact is that collateral knowledge explicit or otherwise is representable as a similarity relation on the conditions of performance. There is thus an intimate relationship between collateral knowledge and the rules that regulate performance. We may illustrate this interrelation with the examples elucidated.

A piece of collateral knowledge that was used in the investigation of the genetic role of nucleic acids was that proteins and nucleic acids are distinct molecular species. Now if we suppose that this is false, then there would be conditions of performance correlated to sentences via the rule R that would be incompatible. For example, consider a sentence "The protein concentration of preparation x is below 0.02%." Certain tests would lead to the acceptance of this sentence while others would lead to its rejection. Let us suppose what is contrary to fact that there existed two bona fide tests one of which measured enzyme activity and the other which measured amino acid content such that in one of the tests the sample

sentence is rejected and on the other test the sentence is accepted. The inferences that concern us here assert that under ascertainable conditions the DNA and not the proteins of certain organisms play certain genetic roles. Hence if there were incompatible tests for determining DNA versus protein content of preparation, then these inferences would be accepted or rejected depending on which one of the pair of incompatible tests were used. No appropriate rule of correlation could be given for the construction of the pragmatic inferences in terms of the set of conditions which contained two incompatible conditions.

Another piece of collateral knowledge used in the same investigations was that the target genetic functions could be recovered in systems which go through transitory periods (e.g., the eclipse period) in which these functions are not manifest. As we have seen, this fact constitutes evidence for the functional separation of the molecular species in question. Let us suppose that this is incorrect, that it is always a DNA protein complex that exhibits the functions in question. We suppose that all the experiments that were designed to separate proteins from DNA and test for the purity of such separation gave the results we have outlined. But also suppose that a DNA protein complex was always formed immediately after the DNA gained entry into the system into which it was introduced and that the genetic functions were the outcome of the cooperative activity of these two molecules. None of the tests described need have failed since they testify to the separability of the DNA from the original proteins and not its separability from some other components with which it may complex after gaining entrance into the reproductive system. Under these hypothetical conditions, it would have been possible to fabricate tests that would vary the type of host and hence vary the type of cooperating molecule that would complex with the input DNA. These tests would yield incompatible results.

As a final example, we may notice that in the second set of inferences described the collateral knowledge of the structure of DNA is of crucial importance. Let us suppose that the Watson-Crick description of DNA was incorrect in that the DNA that is the genetic material of the organisms in

question is a non-linear branching structure. It is possible to imagine situations such that all the experiments described would have yielded the results that they did even with the branching DNA. For example, all of the genetic markers tested could have been on one branch. Under such hypothetical conditions the rules of correlation could have correlated incompatible conditions of performance. Tests carried out with genetic markers on different branches would have been incompatible with tests carried out with genetic markers on the same branch.

These examples tend to show that one of the functions of correct collateral knowledge is to select conditions of performance that are compatible.

REFERENCES

Alloway, J. L., The transformation in vitro of R pneumococcus into S gonus of different specific types by the use of filtered pneumococcus extracts. J. Exptl. Med., 55:91 (1932).

Anderson, T. F., The reaction of bacterial viruses with their host cells. Botan. Rev. 15:464 (1949).

Avery, O. T., MacLeod, C. M. and McCarty, M., Studies on the chemical nature of the substance inducing transformation of pneumococcal types. Induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus Type III. J. Exp. Med. 79:137 (1944).

Basic mechanisms in animal virus biology. Cold Spring Harbor Sym. Quant. Bio., Vol. 27 (1962).

Cavalli-Sforza, L. L., La sessualita nei batteri, Boll. Ist. Sieroterap. Milan., 29:281 (1950).

Chiaraviglio, Lucio, Performance, collateral knowledge and analytic Hypotheses. L'Age de la Science, (1969) (In press).

Delbruck, M. and Bailey, W. T., Induced mutation in bacterial viruses. Cold Spring Harbor Symp. Quant. Biol., 11:33 (1946).

Doermann, A. H., Intracellular growth of bacteriophage. Carnegie Inst. Wash. Yearbook, 47:176 (1948).

Fraenkel-Conrat, H., The Role of the nucleic acid in the reconstruction of active tobacco mosaic virus. J. Am. Chem. Soc. 78:832 (1956).

Fuerst, C. R., Jacob, F., and Wollman, E. L., Determination de liaisons genetiques chez E. coli K 12, a l'aide du radiophosphore. Compt. rend. acad. sci., 243:2162 (1956).

Fuerst, C. R., Jacob, F., and Wollman, E. L. Sur le processus de conjugaison e de recombinaison genetique chez E. coli.VI. Etude du transfert e de la recombinaison a l'aide du radiophosphore. Ann. inst. Pasteur, 101:748 (1961).

Fuerst, C. R. and Stent, G. S., Inactivation of bacteria by decay of incorporated radioactive phosphorus. J. Gen. Physiol., 40:73 (1956).

Garen, A. and Skaar, P. D., Transfer of phosphorus-containing material associated with mating in Escherichia coli. Biochem. et Biophys. Acta, 27:457 (1958).

Gierer, A., and Schramm, G., Infectivity of ribonucleic acid from tobacco mosaic virus. Nature, 177:702 (1956).

Graham, A. F., The fate of the infecting phage particle. Ann. Inst. Pasteur, 84:90 (1953).

Griffith, F., The significance of pneumococcal types. J. Hyg. (Cambridge, England), 27:113 (1928).

Hayes, W., Recombination in Bac. coli K-12: unidirectional transfer of genetic material. Nature, 169:118 (1952).

Hayes, W., The kinetics of the mating process in E. coli. J. Gen. Microbiol., 16:97 (1957).

Hershey, A. D., Mutations of bacteriophage with respect to type of plaque. Genetics, 31:620 (1946).

Hershey, A. D., Spontaneous mutations in bacterial viruses. Cold Spring Harbor Symp. Quant. Biol., 11:67 (1946).

Hershey, A. D., Injection of DNA into cells by phage, in Phage and the Origins of Molecular Biology. editors Cairis, J., Stent, G. S. and Watson, J. D., Cold Spring Harbor Laboratory of Quantitative Biology, Long Island: 1966. pp. 100-108.

Hayes, W., A new approach to the study of kinetics of recombination in E. Coli K 12. Proc. Soc. Gen. Microbiol. J. Gen. Microbiol., 13, II (1955).

Hershey, A. D. and Burgi, E., Molecular homogeneity of the deoxyubonucleic acid of phage T2., J. Mol. Biol., 2:143 (1960).

Hershey, A. D. and Chase M., Independent functions of viral protein and nucleic acid in growth of bacteriophage T2. J. Gen. Physiol. 36:39 (1952).

Hershey, A. D., Kamen, M. D., Kennedy, J. W., and Gest, H., Mortality of bacteriophage containing assimilated radioactive phosphorus. J. Gen. Physiol. 34:305 (1951).

Hodgness, D. S. and Simmons, J. R., Breakage of dg DNA: chemical and genetic characterization of each isolated half-molecule. J. Mol. Biol., 9:411 (1964).

Jacob, F. and Wollman, E. L., Genetic and physical determinations of chromosomal segments in E. coli. Biological replication of macromolecules. Symposia Soc. Exptl. Biol., 12:75 (1958).

Hotchkiss, R. D., Gene, transforming principle, and DNA, in Phage and the origin of molecular biology. editors Cairns, J., Stent G. S., and Watson, J. D., Cold Spring Harbor Laboratory of Quantitative Biology; Long Island: 1966, pp. 180-200.

Kaiser, A. D., The production of phage chromosome fragments and their capacity for genetic transfer. J. Mol. Biol., 4:275 (1962).

Kaiser, A. D. and Hodgness, D. S., The transformation of Escherichia coli. with deoxyribonucleic acid isolated from bacteriophage dg. J. Mol. Biol., 2:392 (1960).

Kaiser, A. D. and Inman, R. B., Cohesion and the biological activity of bacteriophage lambda DNA. J. Mol. Biol., 13:78 (1965).

Lederberg, J. and Tatum, E. L., Novel genotypes in mixed cultures of biochemical mutants of bacteria. Cold Spring Harbor Symp. Quant. Biol., 11:113 (1946).

Lesley, S. M., French, R. C., and Graham, A. F., Breakdown of infecting coli phage by the host cell. Arch. Biochem, 28:149 (1950).

Malφe, O. and Watson, J. D., The transfer of radioactive phosphorus from parental to progeny phage. Proc. Nat. Acad. Sci U.S.A., 37:507 (1951).

Putnam, F. W., and Kozloff, L. M., Biochemical studies of virus production IV. The fate of the infecting virus particle. J. Biol. Chem. 182:243 (1950).

Quine, W. V. O., Word and Object. John Wiley & Sons, Inc., New York and London, 1960 pp. 37-42.

Radding, C. M. and Kaiser, A. D., Gene transfer by broen molecules of DNA: activity of left half molecules. J. Mol. Biol., 7:225 (1963).

Schlesinger, M., The Feulgen reaction of the bacteriophage substance. Nature. 138:508 (1936).

Schlesinger, M., Zwi Frage der chemischen Zusammensetzung des Bakteriophagen. Biochem. Z., 273:306 (1934).

Stent, G. S., Molecular Biology of Bacterial Viruses. W. H. Freeman and Company, San Francisco and London, 1963. p. 114.

Stent, G. S., Mortality due to radioactive phosphorus as an index to bacteriophage development. Cold Spring Harbor Symposia Quant. Biol. 18:255 (1953).

Stent, G. S. and Fuerst, C. R., Genetic and physiological effects of the decay of incorporated radioactive phosphorous in bacterial viruses and bacteria. Advances in Biol. Med. Phys., 7:1 (1960).

Visconti, N. and Delbruck, M., The mechanism of genetic recombination in phage. Genetics, 38:5 (1953).

Watson, J. D., Cold Spring Harbor Laboratory of Quantitative Biology; Long Island: 1966. pp. 180-200.

Watson, J. D. and Crick, F. H. C., A structure for deoxyribose nucleic acid. Nature, 171:737 (1953).

Watson, J. D. and Crick, F. H. C., Genetical implications of the structure of deoxyribonucleic acid. Nature, 171:964 (1953).

Watson, J. D. and Crick, F. H. C., The structure of DNA. Cold Spring Harbor Symp. Quant. Biol., 18:123 (1953).

Wollman, E. L. and Jacob, F., Sur le mecanisme du transfert de materiel genetique au cours de la recombinaison chez E. coli K 12. Compt. rend. acad. sci., 240:2449 (1955).

Wollman, E. L. and Jacob, F., Sur les processus de conjugaison et de recombinaison chez E. coli. V. Le mecanisme du transfert de materiele genetique.

Wollman, E. and Wollman, E., Regeneration des bacteriophages chez le B. megaterium lysogene. C. R. Soc. Biol. 122:190 (1936).